"Implementation of a Route Server for Policy Based Routing across the GIX" Project

Final Report

Tony Bates

RARE
c/o RIPE NCC
Kruislaan 409
1098 SJ Amsterdam
The Netherlands

1. Introduction

The goal of this project is to produce a functioning route server (RS) as specified in "Internet Routing in a Multi Provider, Multi Path Open Environment" by Bates, Karrenberg, Lothberg, Stockman and Terpstra [6][1]. The function of the RS will be to present unified routes to European destinations to routers on the GIX. The results of the project are:

- refining the specifications on routing policy information storage in the RIPE database
- route server software
- software to derive route server configuration information from the RIPE database
- a technical report describing the above

This project requires close coordination with the RIPE NCC for the database related aspects and with the operators of transatlantic links, especially Ebone.

Whilst the aforementioned document contains most of the general concepts and ideas behind the RS some background is given to aid clarity to the report.

2. History and Background

The Internet has evolved from essentially a single top level Administrative Domain (AD) using non hierarchical routing protocols such as the Exterior Gateway Protocol (EGP) [1] for inter-AD routing to an extensive collection of transit ADs with varying policies and requirements.

Due to this added complexity the routing protocols have been improved and the Border Gateway Protocol (BGP) [2] was developed to try to deal for such a complex routing environment. However, today's routing technology and deployed protocols are still based exclusively on destination address which very much limits any flexibility one might want to achieve with routing. With this in mind an increasingly limiting factor is the rapid growth of the Internet in terms of the number of unique paths to any destination rather than just purely the amount of networks themselves. This

1 This document can be retrieved from **ftp.ripe.net:ripe/docs/ripe-docs/ripe-082.ps**

all adds to the difficulty of making the correct routing decision. This can only increase as wide-scale deployment of BGP takes place [3], making for an ever more difficult management task to maintain routing stability leading to large scaling problems.

The Internet Engineering Planning Group (IEPG) [4], the group charged with the technical coordination of global networking infrastructure services, began to foresee these scaling problems. At the IEPG meeting November 1991 in Santa Fe discussion began on the ideas of neutral interconnection and how this could help to achieve needed stability and harmonisation. The general feeling was that if engineered correctly it should be possible to provide a simple point of interconnection for many of the transit service providers allowing them to achieve maximum global connectivity but still making things flexible enough so that special requirements between service providers could be made.

These ideas were discussed and enhanced and at the June 1992 meeting in Tokyo the concept of the Global Internet Exchange (GIX) evolved [5]. The GIX would provide common routing exchange points being independent of any Internet backbone structures. It's primary goals are as follows:

- **Maximise global connectivity**. In other words, make it possible for a new network (generally a service provider) to attach to the GIX and by that have a path to all other connected networks in the global Internet.

- **Provide stable and scalable routing management**. The routing must be extremely stable and any policy-based routing will be left for the networks at the attachment points themselves.

- **Remain flexible**. Make sure that any service provider that attaches to the GIX has the ability to peer with who they want. If there is need for a special policy between any two service providers this can be accommodated.

Along with these goals some key requirements were also needed. The GIX itself must be scalable in terms of both operation and engineering. The management of the GIX must be clearly coordinated as one and at the same time allow the individual participant networks to maintain autonomy.

Within the GIX model are three major components. At the November 1992 informal IPEG meeting in Washington D.C. a paper [6] was presented further refining these components and addressing a possible plan for how these components could be achieved in terms of the European part of the Internet.

The three major components are as follows:

Physical Structure
: A physical media in which participating networks have the ability to attach routers and peer with any other attached router.

Routing Registry (RR)
: A neutral repository used for registering routing policies.

Route Server (RS)
: A source of routing stability coordinating routing information based on the policies set out in the routing registry. A logical meta-router at the physical structure involved in routing

exchanges with the participating attached networks but not taking part in actual packet forwarding.

This report concerns itself with the RS primarily and as will be seen with the aspects of the RR needed for the RS.

3. The Basic Route Server Concept

The Route Server (RS) is meta-router running on a host directly connected to the physical GIX. The task of the RS is to disseminate consistent routing information for all ADs of a region to all ADs connected to the GIX. A RS first peers with all routers that serve the region of the RR to acquire the dynamic routing information for its region. The RS then uses a 'preferred path' database, derived from the RR to filter the dynamic routing information into a consistent routing table for its region. This routing table is then made available via BGP to all routers on the GIX that wish to use it. The RS makes use of the BGP **NEXT HOP** feature so that it will give the correct destination address to be used for a given network.

The RS does not participate in any packet forwarding whatsoever. The only IP traffic is BGP based routing traffic running on top of TCP. There is a clear distinction between routing flow as opposed to packet flow. Figure 1 highlights this distinction graphically.

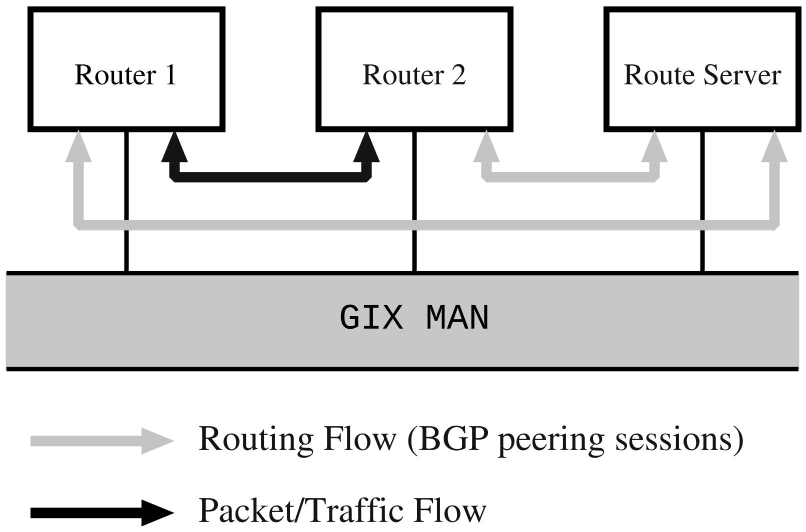

Figure 1: General packet flow versus routing flow

Because BGP maintains consistency of its routing tables when talking to routers within its own AS, this makes it very easy to have more than one RS act as backup. In the long term it would be preferable to have at least two RS's serving Europe.

4. The Routing Registry

As has been stated for the RS to be fully functional all it's routing information should be derived from the European RR.

European IP networking through its natural evolvement has grown into an extremely complex infrastructure and topology which in essence has even today no real single 'coordinated core'. Consequently, there are many differing routing policies used within Europe. The RIPE Routing Working Group was formed to tackle the issues involved in large scale routing collaboration as is needed within Europe. To achieve the needed global connectivity, the network operators and service providers need a way to easily and clearly express the various routing policies and provide a method for problem diagnosis and trouble shooting when problems occur.

The outcome of this is a simple representation syntax for IP routing policies that can be used within the RIPE RR known as **"RIPE-81"**[7][2]. The representation needed to be clear and easily understandable by network operators.

The representation is based upon the assumption that with today's routing technology one network can only have one single routing policy. Similarly, one network can only be in one autonomous system (AS). It gives the ability to depict at the AS level the routing exchanges with other connected AS. The syntax is also rich enough such that it is possible to derive a full graph of routing exchanges within the European Internet. Whilst it is not within the scope of this report to give a tutorial of this representation it should be realised this is key to the RR function and inherently to the function of the RS itself so a brief example is given to show how this routing policy information is abstracted for use by the route server.

As previously noted, a network can only belong to one AS. Hence we have routing policy at network granularity based on AS. It has been argued this assumption cannot be made. However, today we have the problem between "want" and "can" with destination based routing technology and as yet I have not seen a method where "in practice" this assumption doesn't hold true.

The general representation for a network is as follows:

```
inetnum:   192.87.45.0
netname:   RIPE-NCC
descr:     RIPE NCC
descr:     Amsterdam, Netherlands
country:   NL
admin-c:   Daniel Karrenberg
tech-c:    Marten Terpstra
connect:   RIPE NSF
aut-sys:   1104
rev-srv:   ns.ripe.net
changed:   marten@ripe.net 930121
source:    RIPE
```

2 This document can be retrieved from **ftp.ripe.net:ripe/docs/ripe-docs/ripe-081 .ps**

Here there is an AS tag associated with the network database object. If we look at the AS object itself:

```
aut-num:  AS1104
descr:    NIKHEF-H Autonomous system
as-in:    AS1213 100 AS1213
as-in:    AS1913 100 AS1913
as-in:    AS1755 150 ANY
as-out:   AS1213 ANY
as-out:   AS1913 ANY
as-out:   AS1755 AS1104 AS1913 AS1213
tech-c:   Rob Blokzijl
admin-c:  Eric Wassenaar
guardian: as-guardian@nikhef.nl
source:   RIPE
changed:  k13@nikhef.nl 920713
changed:  ripe-dbm@ripe.net 920910
```

The topology of routing exchanges is represented by listing how routing information is exchanged with each neighbouring AS. This is done separately for both incoming and outgoing routing information. In order to provide backup and back door paths a relative cost is associated with incoming routing information. For the purposes of this report, the syntax and semantics are not important. However, we see this information provides both administrative details plus a list of networks, which have a routing policy as depicted with the listed AS neighbours. If we have this information for all the networks in Europe it is possible by a process of recursion to derive the preferred paths from any given network A to network B.

To make this possible a large of amount of effort is needed on two fronts:

Population

Population of the RR is no easy task. However, by providing generalised tools it makes it possible for network operators to see the added benefits of such a generalised routing representation. The most useful tools are the ability to derive router configurations directly from this routing policy information. Table 1 shows the current status of the European RR in terms of routing policy information for known European routed AS's.

Status of AS in RR	Number of AS's	Percentage
In RIPE database with Routing Policy information	42	46.7%
In RIPE database without Routing Policy information	6	6.7%
Not in RIPE database but in NIC/related databases	41	45.6%
Unknown in any databases	1	1%
Total	90	

Table 1: Breakdown of known European routed AS's

Consistency

If the information is to be of use it must be totally reliable. This is the primary task of the RR. Consistency and elimination of any conflicts must be maintained at all time.

RIPE-81 has been widely accepted by the European community as well as recognised world-wide as one of primary methods of registering IP routing policy information. This is a major achievement, due in large part to the RS project.

5. RR/RS Interaction

The key to the success of the RS is the ability to extract the routing policy information from the RR. This is done by a process of derivation of the RR information until a point whereby it is possible to determine which is the preferred route from the GIX for any European destination. Figure 2 depicts the general RR/RS Interaction concept. A routing policy template is sent into the RR and after parsing and consistency checking the relevant part of the RS configuration is generated. The RS then uses this information to inform the routers at the GIX which is the preferred route for a European destination. As stated above we still do not have all the information needed in for the **full** concept to be used so a sub-set of this is used. This will be explained in a later section.

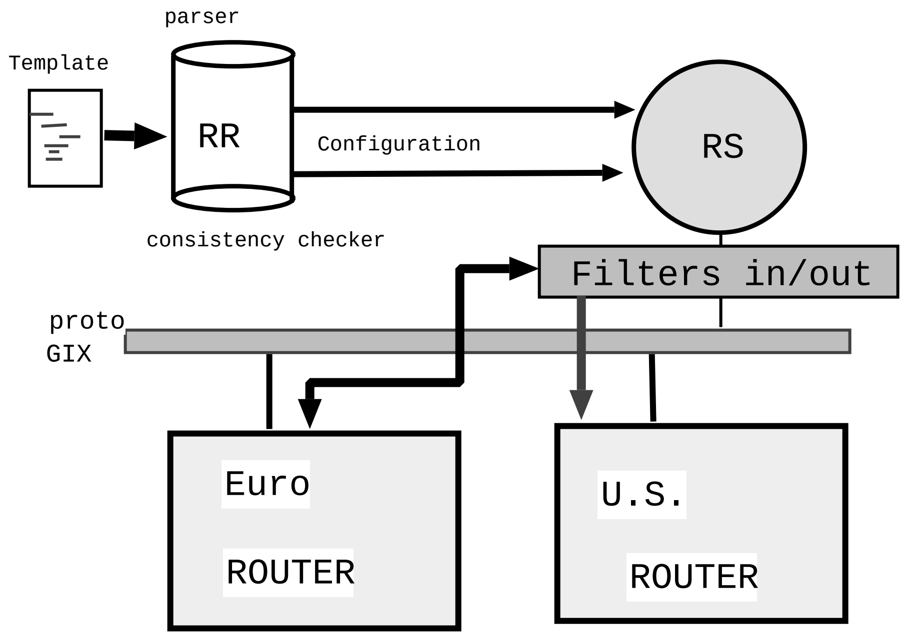

Figure 2: RS/RR Interaction

6. RIPE Route Server Basic Set-Up

With the help of SUNET, Alternet and Sprint the European RS was attached to the GIX in early March 1993. The Hardware is based on a SUN IPC with 32 Mbytes of memory and 400 Mbytes of hard-disk. This was deemed to be more than adequate. For those that way inclined the machine is known as **europe-rs1.ripe.net** and is located within the Sprint POP in Washington D.C.

The RS software itself is based on the *GateD* software from the Gatedaemon Project at Cornell University. It only makes use of the BGP function within GateD. By making use of the information derived from the RIPE RR, **network** based filter lists are created to derive the preferred paths for European networks at the GIX.

Network operators in Europe send in their routing policy information to the RIPE RR and from this information a daily run of the database produces the needed RS configuration which is dynamically loaded into the RS without any interruption to the RS service.

Referring to Figure 2 we see how this information is disseminated. The key point to note is that there are no in-bound routes accepted from the U.S. router whatsoever, just the dissemination of European routes from the RS. In practice, the distinction will not always be as clear cut at the conceptual diagram but the principal holds no matter how European networks are connected to service providers at the GIX.

At the present time this is only done on an experimental basis whilst the software is proven to be stable and all the RR information is gathered. Some test networks are used in a "production" manner and no problems have been encountered so far.

6.1. BGP Problems ?

Early on in the deployment a large amount of testing was done with the BGP implementation in GateD. It can handle BGP 2 and BGP 3. It has been tested against the following implementations:

1) Cisco (various versions) — running both BGP2 and BGP3

2) ANS RS/6000 Router ‘rcp_routed’ — running BGP 2

Several EBGP based tests were carried out. Specifically to see if consistency of AS path information was maintained.

The GateD BGP implementation is highly configurable and allows you to do both AS-path and network based filtering with a powerful redistribution mechanism. However, the syntax is somewhat "non-user friendly" and open to human error so it is important to always pre-parse (the "-C" option in GateD) any configuration before trying to update the configuration.

6.1.1. The "NEXT_HOP" Bugs

Unfortunately, I found bugs in both the cisco and rcp_routed implementations of BGP. A little worrying as a correct NEXT_HOP is the foundation of the RS concept. These have been subsequently fixed (or re-engineered in one case).

These problems generally occur with the use of secondary addresses or being multi-homed and passing a network advertisement with a NEXT_HOP of the multi-homed network which is not directly connected to the GIX.

For example, if we look at Figure 3.

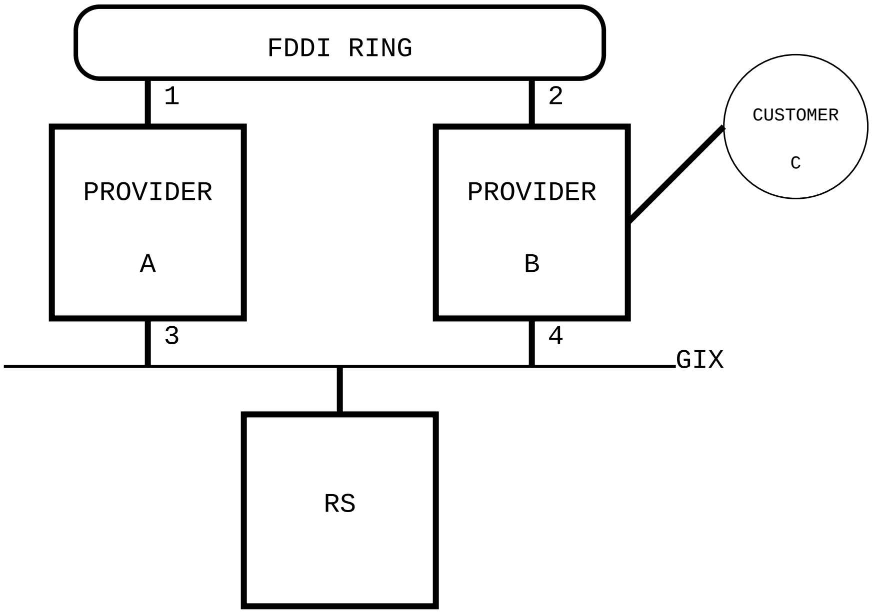

Figure 3: Example of "BAD NEXT_HOP" problem

Provider B has a route to customer C which is advertised to Provider A both via the GIX and a private connection (in this case an FDDI ring). The private connection is preferred (for obvious reasons) and provider A has a route to customer C of interface 2. However, when that route is advertised to the RS it advertises a BGP NEXT_HOP of interface 2. This is a bogus route as the RS cannot talk directly to the interface. The correct thing to do in this case is close the BGP connection (as per the specification). However, a local modification was made to install a route of interface 3 in the RS. This modification will be in future releases of GateD to make it configurable as to what to do with a bad NEXT_HOP.

6.2. GateD Performance Issues

The performance needed can de divided into two areas:

1) BGP performance

   The BGP performance has been found to be more that adequate. The RS currently has 7 EBGP peers (see later section for actual set-up. It processes approximately 6000 protocol changes (i.e. a change, add or release of of BGP route) in 20 minutes or around 5 a second. These measurements are done by the examining the GateD logs. The logging mechanism I use is to have 7 megabytes of log space which are rotated into 1 megabyte files. These files rotate approximately every 30mins which gives an indication of how active the routing is.

2) Policy List performance

   As stated the RS currently uses network based policy lists (the generation of which is given in a later section). Currently, this list is loaded in once a day. The size of this list currently amounts to 7000 networks (many networks have more than more path from the GIX and have a unique entry per BGP peer they connect to).

The average time to re-read the policy list is 4-5 minutes. During this time GateD will stop listening to its open sockets while it linear scans and re-builds its internal policy list. It is possible to run out of buffers on the sockets or to reach bgp_traffic_timeout from the remote peers. At this point the peer will be reset. This is a serious performance problem due to the linear sort of the policy list. Some performance gains can be made by building the network list to strip any trailing zeros for the networks and sorting them in reverse order. This gains a little and depending on the states of the BGP peers so the RS does not always drop the peer on configuration re-read.

However, a "radix sort" algorithm will soon be put into GateD which will effectively remove this performance overhead. A 7000 network list will mean approximately 32 instructions rather than current 7000. This will be available in the next release.

6.3. Operational Issues.

The RS is operated by the RIPE NCC remotely from Amsterdam. This is not without some operational issues. However, a direct **clone** machine is installed in Amsterdam. All development and testing is done on this machine and when shown to be stable transferred to the production machine. Good out of band access is provided and local facilities management means no real problems exist with the machine being located remotely.

A deployment plan[3] was devised and is currently *on schedule*. Much of the early trialing was based upon testing and interaction with other service providers at the GIX. A certain amount of *trust building* is needed and it is clear the cooperation with the service providers at the GIX is needed and appreciated.

A loan cisco has also been provided courtesy of SURFnet to test the RS software locally and provide a mechanism of creating a variety of NEXT_HOP scenarios.

6.4. How is the RS configuration built today ?

As has been pointed out several times throughout this report, we still do not have all the RR information to completely generate the configuration transparently. However, an interim or more precisely a "kludge" mechanism is in place. The "kludge" uses a simple configuration where the peer (at the GIX) is given an associated TAG and an AS (remember in RIPE-81, an AS means a list of **unique** networks) has a preference associated with it. It then gets the list of networks from the RR wherever it can and if currently it can't it derives them from a *well connected router* within Europe. To make sure things are up to date a program is used to monitor any new AS's seen in routing tables so the "kludge" builder is up to date.

The AS then has a preference of TAGS associated with it to generate the network preference lists.

Here is part of an auto-generated "kludge" configuration which is used to generate the GateD network list.

---

3 See Appendix A.

```
#
#Kludge config generated at Jul 25 13:38:09
#
AS1103:SURFnet autonomous system:R#P:NSS:ICM
#
AS1104:NIKHEF-H:R#P:NSS:ICM
#
AS1109:UNI-SALZBURG ACOnet:N#W:NSS:ICM
#
AS1110:UNI-INNSBR ACOnet:N#W:NSS:ICM
#
AS1835:DENet Danish University Network:R#P:ICM:NSS
#
AS1849:PIPEX, Public IP EXchange:R#P:ALTER:ICM:NSS
#
AS1853:ACOnet Backbone:R#P:NSS:ICM
#
```

Here we see some ASes. The names comes from the RR as does the third field which is an internal representation telling the builder whether the AS can be derived from the RR or not. The rest of the fields represent TAGS for peers.

ICM

The ICM router at the GIX.

NSS

The ANS NSFnet router at the GIX.

ALTER

The AlterNet router at the GIX.

There are other TAGS for the other peers and this is just an example.

The ordering is used to determine which is the preferred route from the GIX. From this we end with a network list with GateD preferences which is used by the RS.

While this approach is a kludge in that some of the information has to derived from dynamic sources (i.e. routers) rather than a central repository it does test all the principles of the RS as well as providing a basis for the performance potential of the RS.

7. Current Configuration

Currently, the RS is peering with almost all the routers at the GIX. Figure 4 show the actual details including the AS numbers of the GIX partners as well as where they are located.

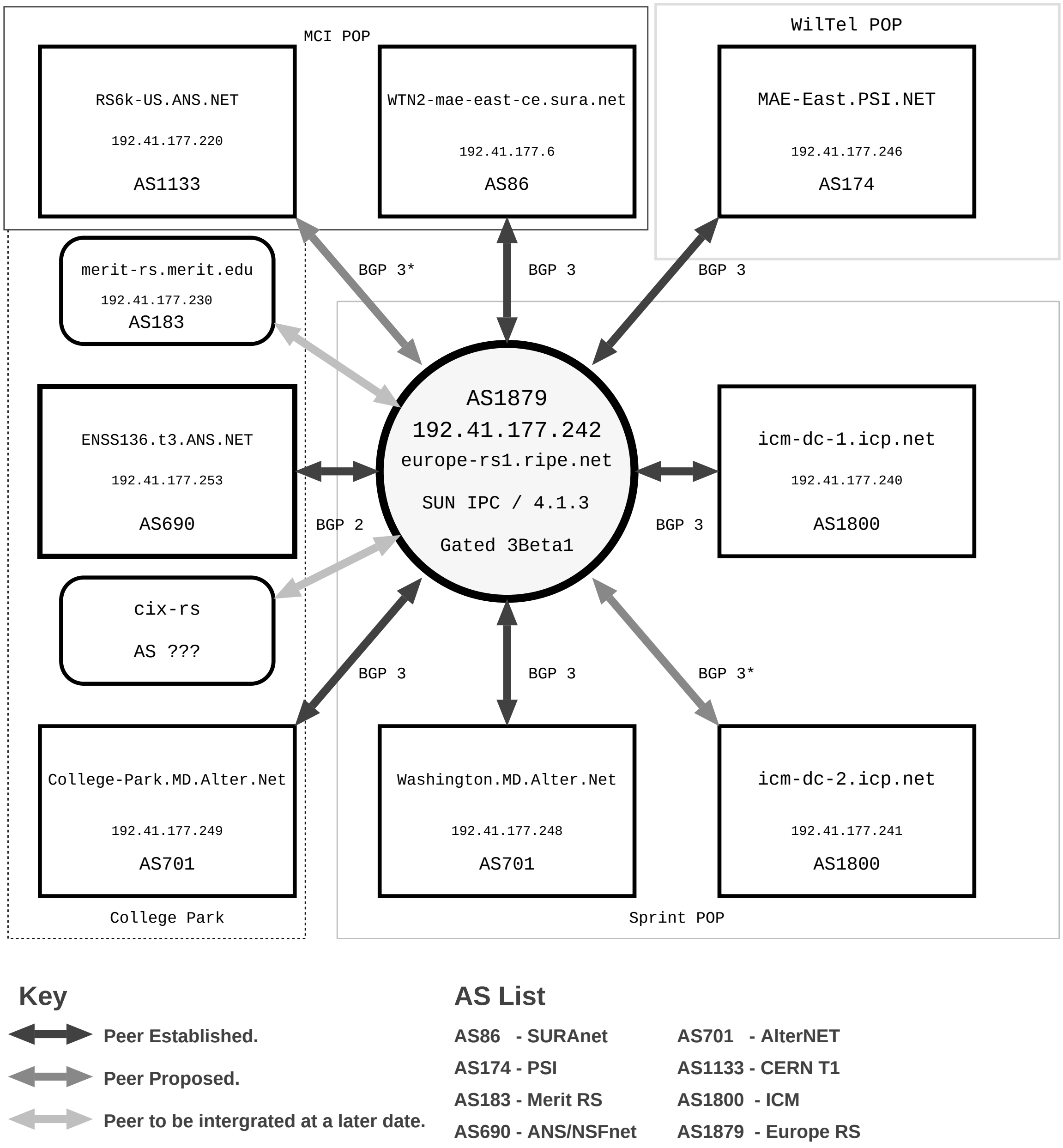

Figure 4: Current configuration of the RS

The peer still to be integrated is the new RS6000 router in use for the CERN T1. This awaits ANS to configure their peer. A Sprintlink (rather than ICM) router will be attached to the GIX and peering can begin straight away. Currently, only ENSS 136 accepts routes (some test networks) from the RS.

7.1. How Soon can the RS be used for Operations

At the recent IETF in Amsterdam a meeting was held between the operators for the service providers at the GIX and a plan to integrate the European RS was devised. The plan is somewhat in line with the original deployment plan with one caveat.

It is clear that the current routing situation is becoming increasingly difficult to maintain and the operators would like to see something to make their lives easier. The RS project is committed to using "network" based list until I can be sure we can trust AS-paths for real-time operation. However, in the interim the ICM router has been elected to provide AS-path based route redistribution of European networks based on the "kludge" configuration which will be maintained by the RIPE RR. As soon as the information is in the RIPE RR the operators are happy to switch to the European RS as the primary service point for their route determination with ICM as backup.

Unfortunately, the minutes of this meeting were not available in time for this report but this represents a major breakthrough in terms of the RS project. We appear to have the trust of the operators both in terms of the routing policy information we store in the RIPE RR and in the operation of the RS itself.

8. Other RS projects

There are currently two other RS projects in progress or about to start.

1) The "Merit RS"

   The Merit RS is also deployed at the GIX and Merit have stated an intention to handle all routes not serviced by the European RS if needed. Their current thinking is to base RR information on the NSFnet Policy Routing database in use for the current NSFnet backbone. This does not have some of the advantages of the RIPE-81 ideas in that it the NSFnet backbone in intrinsicly tied into the database. However, work is underway to use a similar approach as used in RIPE-81 with the recognition of an "Originating AS" attribute in their database.

2) The "CIX association RS"

   The CIX association plan to use the RIPE RR software and tools to create a CIX database (including the RR information) and a CIX RS. Much of their deployment is dependent on the database software currently under revision at the RIPE NCC.

A large amount of collaboration takes place between the three projects under the discussion list of **rs-imp@merit.edu**. It is hoped in the future that the RS's will in fact be able to act as backup for each other. For this to occur, full consistency of which RR maintains routes for which networks needs to be ratified.

9. How to get the RR populated ?

Whilst the basic components of the RS function, the missing part is the rest of the RR information. As stated, the way to do this is by the introduction of tools and convincing the European operators in the benefits of making their routing policy information available. This is in practical terms the second phase of the RS project. A list of useful tools have been identified as follows:

prcheck — A tool to check the consistency of routing policies stored in the routing registry. This tool will flag if two neighbouring network operators specify conflicting or inconsistent routing information exchanges with each other and also detect global inconsistencies where possible.

prpath	Extract all (AS-)paths between two networks which are allowed by routing policy from the routing registry.
prconn	Display the connectivity a given network has according to current policies. This will of course also be able to find the set of networks a given network can not reach.
prtraceroute	A version of the existing traceroute tool which will be able to display whether a route in use is allowed by policy and where deviations from policy occur.
prconfig	A router configuration generator. This will generate configuration files for specific routers from the routing registry and additional information supplied by the user.

The development of these tools is outside the scope of the RS project but a proposal has been produced for a separately funded project known as "PRIDE - Policy based Routing Implementation, Deployment in Europe" [8].

10. The Future

There are many who see the RS evolving into an even more specialised routing engine than it already is. One which can handle multiple routing tables depending on which peers need the routing information. However, until we see whether we can make a "distributed RS model" work it is unclear if this effort is needed or possible. The key component (as with all operations in fact) is easy access and derivability of the routing policy information. Until we have this it seems unlikely we will see such a full blown RS under development.

Much of this depends on the ability to make use of the IP source address as part of the routing decision process. There are proposals currently under consideration within the IETF that may help and at the very least need to be tracked. Specifically, the work on IDRP [9] and SDRP [10].

The RS must be able to deal with CIDR [11] based addresses. The current plan is to upgrade to BGP4 and use CIDR in line with the current plans for CIDR deployment within the Internet [12]. For this to be of real use there are many outstanding issues regarding registration of CIDR based routing policy that will need to be solved. However, these are beyond the scope of this project.

11. Conclusion

The RS project has been a success. We have a deployed a route server at the GIX and put the building blocks in place for the formation of a European IP routing registry. The goals of the project have been met and have helped to progress other RS initiatives with the Internet community. Europe has shown a lead role in the development of a generic routing policy information representation which is receiving alot of attention within the community.

The PRIDE project should help to build on the work currently in place from the RS project. Whilst not directly in the mandate of PRIDE, it is hoped that the RS progress can continue at the rapid rate in which it has up to now.

The future of the RS is still a little unclear in terms of whether something more complex will be needed in the long run. This can only be determined when more experience is gained with the RS in operation and with the introduction of other RS's at the GIX.

12. References

[1] Lougheed, K., Rekhter, Y., "A Border Gateway Protocol 3 (BGP-3)", RFC1267, October 1991.

[2] Mills, D., "Exterior Gateway Protocol Formal Specification", RFC 904, April 1984.

[3] Yu, J., BGP Deployment and Application (bgpdepl) Charter", bgpdepl-charter.txt (Working document), April 1993.

[4] Huston, G. "Intercontinental Engineering and Planning Group Inet'92 Report", Proceeding of INET'92, pp 239-247, June 1992.

[5] Almes,G., Lothberg, P., Ford, P., "Proposal for Global Internet connectivity IEPG Working Document", June 1992.

[6] Bates, T., Karrenberg, D., Lothberg, P., Stockman, B., Terpstra, M., "Internet Routing in a Multi Provider, Multi Path Open Environment", RIPE-82", March 1993.

[7] Bates, T., Jouanigot, J-M., Karrenberg, D., Lothberg, P., Terpstra, M., "Representation of IP Routing Policies in the RIPE Database", RIPE-81, February 1993.

[8] Karrenberg, D., "P R I D E - Policy based Routing Implementation, Deployment in Europe, A Project Proposal", July 1993.

[9] Steenstrup, M., "Inter-Domain Policy Routing Protocol Specification: Version 1", INTERNET-DRAFT, May 1992.

[10] Estrin, D., Li, T., Rekhter, Y., "Source Demand Routing: Packet Format and Forwarding Specification (Version 1)", INTERNET-DRAFT, March 1993.

[11] Fuller, V., Li, T., Yu, J., Varadhan, K., "Supernetting: an Address Assignment and Aggregation Strategy", RFC-1338, June 1992.

[12] Topolcic, C., "Status of CIDR Deployment in the Internet", RFC-1467, August 1993.

13. Appendix A

European Route Server Implementation Plan

Tony Bates
RARE
Amsterdam

(Version 0.2)
WORK IN PROGRESS

Introduction

The plan outlined is aimed at piloting the concepts introduced in the IEPG proposal for a European Route Server [1]. One should be familiar with this proposal. The main issue of the pilot is to investigate the feasibility and work required in achieving 'routing registry/route server' interaction rather than transport aspects of the GIX. An implementation plan including some loose timescales is discussed and tabled below.

Environment

The Project falls into two areas of interest and some detail of the environment for these areas of the project is deemed useful at this point;

(1) **Routing Registry**

One of the critical pieces of the route server is to make sure that you guarantee correct and consistent routing policy information for all the IP networks that you wish to provide a route server service for. What this essentially means is that the creation of a database of all European IP networks containing a method of documenting their chosen routing policy is needed. This is understandably no small task, especially when the scale of coverage is one as large as Europe. However, a strong initiative has already begun to make this database a reality. At the RIPE meeting in Prague in January a paper [2] was presented which discussed some new objects within the RIPE database which make it possible to document and express the routing policy for any given IP network within Europe. Through the channels of the RIPE routing working group, general consensus has already been reached on the new database objects and work is under way on a refinement of the text and some enhancement of the objects so that work can begin on piloting the new objects. The proposal itself is essentially more generic than just for use with the route server. However, it contains all the information needed for the European part of IP routing policy information.

Along with the database itself, a large amount of effort is needed in actually collecting the routing information for the database. One of the largest and perhaps most difficult parts of the project is educating and communicating to the network providers the merits of such an exercise of

[1] "Internet Routing in a Multi Provider, Multi Path Open Environment", "ripe-82", Tony Bates, Daniel Karrenberg, Peter Lothberg, Bernhard Stockman and Marten Terpstra. February 1993.
[2] "Representation of IP Routing Policies in the RIPE Database", "ripe-81", Tony Bates, Jean-Michel Jouanigot, Daniel Karrenberg, Peter Lothberg and Marten Terpstra. WORK IN PROGRESS, February 1993.

information provision on their part. Working at the RIPE NCC allows this to be done in the most neutral way possible. This can only be done by convincing the network providers as to how it will benefit them and the global Internet now and in the future. As essential part of the project is to produce a list of tools.

(2) **Proto-GIX**

Under the auspices of the IEPG, we are fortunate enough to have the potential to pilot a basic route server using a‘GIX like’ or proto-GIX infrastructure that encompasses many of the international service providers today. This infrastructure is a Metropolitan Area based network (MAN) within the Washington D.C. area known as the Metropolitan Area Ethernet — East (MAE-EAST). This provides a basic 10 Mbps MAN that is conveniently located in the sense that several of todays service providers have or will have connections to this MAN. It is also fortunate for Europe to pilot its RS in the US as it provides the ability to present a full consistent view of European routes to the US as many of the current European transatlantic links terminate at or with service providers with connections in the D.C. area.

The practicalities of this mean that by making use of out-of-band access it will be possible to develop, test and examine software for the route server in a ‘real life’ GIX type environment. Initially there need be no service reliance on the route server itself. A service provider can peer with the route server and still continue to peer with the other providers routers on the GIX. By making use of metrics or some form of administrative distance it will be possible for the service provider to examine the routes imported form the route server without making any use of the information for production traffic itself. More details of this aspect are given Section 3.

Implementation Plan

The plan outlined in this section is somewhat loose and certainly will be changed and modified in the light of experience gained in the project. The plan below concerns itself mainly with the route server aspects of the project directly rather than the secondary aspects associated with the route server including some of the items outlined in Section 2.

The time-scales noted should are also to give a basic time-line for the various aspects of the plan rather than anything that will be ‘cast in stone’.

—— February, 1993 — WEEKS 5 - 8 [RR] ——

1) Define new objects and definitions required in the RIPE database.

Essentially produce a paper to propose and enhance the current database objects so we have the needed information to populate the RIPE database. This calls for some clarification of the current thinking of routing and routing policy representation within the RIPE database.

2) Produce concept and procedure documents for service providers.

With the introduction of the new objects within the database and there foreseen use for routing policy, a usage/impact document is required for the service providers. Without the complete understanding and cooperation of the service providers the information required will not be forthcoming.

3) Gather information and feedback from service providers.

As part of this action, a pilot will take place with a small group of European service providers. Within the pilot the following development aspects will be covered.

3.1 **Modification of Database software**

The RIPE database will need to be modified to take into account the new proposed database objects.

3.2 **Develop a transition method for existing routing pilot.**

A transition plan is required to convert existing routing pilot [3] information into the new proposed database format.

3.3 **Produce Configuration based Tools**

Tools are needed to derived router configurations from the routing policies expressed within the RIPE database.

3.4 **Enhance the current database update mechanism for routing policy tags.**

An enhanced and more usable update procedure is required.

3.5 **Database generation scripts.**

Create scripts to generate existing network information data within the the database into the new formats.

—— **March, 1993** — WEEK 9 [RS] ——

4) Install Route Server on Proto-GIX.

Physically attach the route server to the GIX itself. Verify and modify OS environment. The Route Server machine itself with be a SUN IPC with vanilla 'gated' software. This should be run for at least one week to examine stability, basic operational aspects before any routing is configured or tested.

Start to load in trial configurations, including routing filter lists derived from the database.

—— **March, 1993** — WEEK 10 [RS/RR] ——

5) Establish inbound peer sessions to Routers attached to the GIX.

In this initial period, only import routes from routers attached to examine and verify the routing information. Do not export any routes whatsoever.

Examine stability of route server routing software.

Start to verify the exported routing information in terms of the routing policy information derived from the RIPE database.

6) Establish peer session with the NORTH AMERICAN route server.

---

[3] There is an existing 'routing' pilot using the database objects outlined in 'ripe-60'.

Exchange European routes to the north American router and examine consistency of information. Try to eliminate areas of duplication. Investigate convergence issues involved between two route servers.

—— **March, 1993** — WEEK 11 [RS/RR] ——

7) Export routes to "selected" service provider router.

Begin exporting routes to a "selected" service provider router. By making use of either metrics in the route server or weighting the routing announcements to the provider router, evaluate the routing information provided to the provider router without making use of the information for any forward purposes.

Work with the provider to verify and check consistency of exported information.

—— **March, 1993** — WEEK 12 [RS/RR] ——

8) Increase list of routers to export routes to.

Add more ‘export’ neighbors. Examine and test the route server routing software under load.

9) Use the derived routes from route server for operational use.

With a "selected" service provider, adjust the metrics accordingly to select the routes exported from the route server in preference from the routes derived from other service provider routers, whilst continuing to maintain the existing peerings as back-up. Produce a ‘simple’ migration document from this to be used with other service providers.

—— **March, 1993** — WEEK 13 [IEPG] ——

10) Summarise pilot to date at ‘informal’ IEPG, Columbus.

Present pilot results. Discuss with other service providers and North American route server project the results and needed enhancement.

Investigate the possibility of establishing an Asia-Pacific route server at the GIX.

—— **April, 1993,** — WEEKS 14-16 [RS/RR] ——

11) Draft Operational Procedures for a production route server.

In collaboration with the IEPG, service providers and the North American pilot begin to develop operational procedures and requirements for a production route server.

Encourage more use of the route server for production based route derivation.

12) Produce Intermediate Report for circulation at Amsterdam IEPG.

Produce a summary of pilot results and produce implementation plan for second half of project.

Discuss and develop project plan for second half the project to present results at the Amsterdam IETF/IEPG in WEEK 28.